DOCUMENT RESUME

ED 170 687 CS 004 513

AUTHOR Stone, David E.
TITLE The Semantic Content of Text, Pictures, and Events. Research Report No. 1, Series B.
INSTITUTION State Univ. of New York, Ithaca. Coll. of Agriculture and Life Sciences at Cornell Univ.
PUB DATE 25 May 78
NOTE 24p.; Paper presented at the Symposium on Research in Graphic Communication (Oxford, England, May 25, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Processes; Diagrams; *Illustrations; *Reading Comprehension; *Reading Processes; *Reading Research; *Visual Aids

ABSTRACT

The degree to which printed text and graphics complement each other to facilitate reading comprehension was studied with the following four research tools: a system for analyzing text and graphics to reveal aspects of their form and content; procedures to permit analysis of both verbal and nonverbal recall measures of comprehension, allowing comparisons to be made with the original passage; eye track programs written to display text and graphics, collect data on readers' eye movements, and analyze the collected data; and a computer data analysis system. Research results indicated that certain specific types of information appeared to be better understood when presented in a particular form of text, graphics, or combinations of the two. The following specific results emerged: spatial or locative information is more completely and accurately understood when presented graphically than in text when the semantic content of each form is equivalent; information must be written and read in linear sequence; locative information presented in text form leads to comprehension difficulties; and sequential information is more effectively communicated in text than in illustrations. (Illustrations are included.) (DF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED170687

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Semantic Content of Text, Pictures, and Events

by

David E. Stone, Ph.D.

Research Report No. 1 Series B

A paper presented at the
Symposium on Research in Graphic Communication
The Open University Reading Research Group
Oxford, England
May 25, 1978

Reading and Learning Series
Research Report

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
David E. Stone

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Reading Research Group
Department of Education
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University
Ithaca, New York 14853

CS004513

The research reported here was supported by Hatch Grant No. 406, Marvin D. Glock, Ph.D., Project Director and by SUNY Research Foundation Grant No. 7422A, David E. Stone, Ph.D., Project Director.

The degree to which printed text and graphics complement each other to facilitate reading comprehension is a problem of increasing interest to American psychologists and educators. The work that I shall be presenting today has been conducted jointly by reading research groups at Cornell University (under the direction of Professor Marvin D. Glock) and at the Child Research and Study Center of the University at Albany, one of the research centers of the State University of New York.

The principle objective of our research has been to learn more about the reading process, which we believe properly includes the study of the process involved in understanding graphic as well as written language. It is hoped that the major by-products of such research will include a better understanding of the reading process as well as a body of knowledge, based on research, that will make possible the improved design of reading materials which employ text and graphics.

We have taken the position that we will learn the most by asking questions about the processes involved in comprehending text and graphics, the nature of the mental representation of their meaning, the basis on which the reader selects some information to retain rather than other, how information in the passages is integrated with related prior knowledge and how it is affected by the acquisition of later knowledge, how stored information is later retrieved, the nature of the processes involved in making use of that knowledge for the purposes of the task used in assessment, etc.

In our work we have used both free recall (that is, the procedure of asking people to write what they remember after reading) and video-tape records of an assembly task to provide the rich sources of data we believe

necessary to examine these questions. We have also begun to examine differential scanning patterns (eye-movements) for clues to the nature of the reading process.

This program of research has necessitated the use of several, rather sophisticated, tools. First, we have employed a system for analyzing the text and graphics to reveal aspects of their form and content. (Such a system must permit the evaluation of these aspects of both text and graphics in comparable terms.) Second, procedures had to be developed which would permit analysis of both verbal and non-verbal recall measures of comprehension allowing comparisons to be made with the original passage. Such procedures must permit the analysis and comparisons to be made in equivalent terms. Third, eye-track programs had to be written to display text and graphics, collect data on eye-movements of readers, and to analyze the collected data. Finally, it has been necessary to employ a computer system to manage the large quantities of data produced by such research. I will now examine each of these problems and then present some of our preliminary results.

A System for Analysis of Text and Graphics

If we wish to determine what information is retained from a passage by a reader, the first step is to determine what information is expressed in the text and graphics that comprise the passage. As I have also indicated it is necessary to measure such information contained in text and graphics in equivalent terms so that the relative contribution of each can be assessed, at least to some degree.

It has become common in recent studies of the reading process employing text to make a distinction between form and content, assuming that the same content can be present in text passages having different form (McConkie,

1977). It has been our position that that same content can be expressed in a variety of forms of text and/or graphics. Our current research has been directed towards learning more about the effects of manipulations of content and form of text and graphics on what readers understand and remember.

Let us consider the problem of determining the content of a passage including text and graphics using examples from McConkie (1977) originally designed to deal exclusively with text.

A passage of text or a graphic can be viewed as presenting a number of different concepts generally classifiable as objects, actions, or attributes. The passage of text expresses relationships among these concepts while an illustration represents such relations graphically. The passage or graphic can therefore be represented as a form of network structure, with the concepts being the points or nodes in the structure, and the relations being represented by the lines between the nodes. These lines can be labeled to indicate the type of relationship between the concepts. (For example, an object, - John - may be either the one who performs an action, - hit - or the one who is the object of the action, the one being hit. The label on the line is therefore either object or agent to make the necessary distinction.) (McConkie, 1977)

In order to determine the content of a text or graphic that is transformed by scorers into a series of propositions expressed as a network structure, let us consider another example from McConkie (1977) - John hit the red ball - is represented by two propositions: "The ball was red"

P1: (:Ball)--att-->(red)

att = attribute

and - "John hit the ball" -

P2: (:John)--agt---(hit)--obj---(:Ball)

agt = agent, obj = object

Words representing the concepts mentioned are placed in parentheses, and the labels on the relationships are embedded in the arrows which represent these relationships. Each proposition is numbered.

Figure 1 shows a sample of a procedural text used in one of our studies. It is simply a portion of the directions for the assembly of a model.

-- Insert Figure 1 about here --

Figure 2 shows the content of this passage as represented in this

-- Insert Figure 2 about here --

scoring system. I will forego an analysis of each proposition, but will instead refer to Figure 3 which shows a series of illustrations used with

--Insert Figure 3 about here --

text directions. The two frames that are checked correspond to the text directions I have just shown you. Figure 4 shows the representations of the content of these two illustrations in the scoring system. The first step in such an analysis is to ask several trained scorers to each write a text to depict the content of the illustration. One may reasonably inquire at this point why we have chosen to represent the content of the graphic in text -- and not the reverse. The answer is simply that to do so would frequently require the use of animations while the representation of graphic content in text is a more practical procedure.

Another question may also be raised here. Why not develop a scoring system that permits representation of the content of text and graphics in equivalent terms without going through any intermediary step? In fact, we

could do this but have found that, at this point at least, that our scorers are not sufficiently fluent in the systems used so as to apply it directly to non-verbal representations.

For these reasons we begin with the translation of graphic content into text form. For the procedural passage presented here scorers refer to a standard lexicon for the names of objects and features of objects depicted in the graphic. The relations among these objects (such as their orientation in space, number, location, etc.) are expressed by the scorers in whatever terms they find most comfortable.

Once several text representations of the content of the graphic have been prepared they and the text portion of the passage are subjected to a further independent analysis using a scoring system which transforms the content of each into a propositional network structure, as shown in Figure 4.

-- Insert Figure 4 about here --

In looking at Figure 4 you will notice that after several propositions the words -- NOT PRESENT -- are shown. The numbering of the propositions in this case is the same as for the text. It should be clear then that several of the propositions expressed in the text are not present in the illustration which accompnay the text. It may also be observed that a new proposition is present here that was not present in the text.

Such information is just the sort required if we are to consider how the content of text and graphic complement each other. The advantage of this scoring system for us has been that it forces us to examine gross differences between the semantic content of text and graphic, such as these. It also leads us to conduct a particularly fine-grain analysis of the logical and semantic content of text and graphics. The scorer's product is as close to a complete and objective specifications of content as we believe necessary for our purposes. It's ability to express such content of both text

and graphics in comparable terms is crucial for such comparisons.

One important drawback to this system is that the analysis of even the briefest text or simplest graphic yields a vast number of propositions. It has, therefore, been important to develop computer programs to make the comparisons between the scorer's representations of the semantic content of text and graphics. The first of our programs for this purpose is now being developed. We hope someday that the entire scoring process can be done by computer, but that is a long way off. However, the scoring system itself is the first of our four major tools for exploring this research area.

Our second requirement was for a set of procedures which would permit analysis of both verbal and non-verbal recall measures of comprehension allowing comparisons to be made with the original passage. This use of free recall measures (asking people to write down what they remember after reading) along with the use of non-verbal measures (such as assembling a model, or operating equipment) seems to have potential for helping us learn more about the manner in which knowledge is stored in memory. Differences in the verbal and non-verbal recalls of the same content may be useful in this regard.

In any case, researchers are now using both these means of assessing comprehension and therefore it has been necessary in our work to find a way to apply the scoring system I have described to the recalls. Such an application permits a precise comparison of the differences between what has been presented to the reader and what he remembers. The analysis of the free written recalls is exactly the same as for the original text. The non-verbal recalls require an additional step.

In our work, non-verbal recalls are recorded on videotape and then subjected to a second-by-second review. Scorers generate written descrip-

tions of recall in much the same way they score graphics. That is, they write descriptions and then analyze them. These representations are then compared for a measure of reliability. Upon completion of this process, comparisons between the original text and/or graphic can be made with the verbal and/or non-verbal content of the original materials. This then satisfies our second express need in this area of research.

Once a set of free recalls of a passage has been analyzed and scored, the resulting data can be examined in several different ways. For example, it is possible to examine the frequency of individual elements or propositions, the sequence in which propositions are recalled, the ways in which recalls differ from the original materials, with additional or deleted information and so forth.

Such data relates to questions such as what structural characteristics of form or content affect the recall of elements of the text or graphics or what meaning readers seem to extract from a portion of text or graphic.

The third approach in our research has been to use records of reader's eye-movements as clues to the nature of the processes involved in the extraction of meaning from text and graphics.

The Eye-track System

Our interest in the use of eye-movements to help understand the way people read text and graphics is based on evidence provided by a number of researchers that eye-movements are a valid indication of cognitive (mental) events. Contemporary researchers taking this perspective include George McConkie of Cornell, Keith Rainer, University of Rochester, and Geoffrey Loftus, University of Washington.

Our preliminary investigations using eye-movements have employed videotapes to provide gross measures of the amounts of time subjects spend

looking at portions of text and graphics, the frequency with which they refer to text or graphics and the sequence in which they do so. Figure 5 depicts a laboratory set up for this purpose.

-- Insert Figure 5 about here --

Our method for collecting such data has been to present individual readers with slides of graphics and slides of text shown on two separate screens. With adequate separation of the screen, an accurate indication of the reader's eye-movements can be obtained by a camera positioned adjacent to the screen. (For some studies we have used a second camera to record performance measures of reading comprehension.)

These tapes are then subjected to a second-by-second analysis. The theoretical framework within which we have viewed these data has been based on work by Loftus (1976) on picture recognition. Loftus views the process of encoding information in graphics as consisting of three stages.

Stage 1 -- Decision about where to look in the picture

Stage 2 -- Once a particular area has been fixated, information from this area must be extracted and processed during the fixation, and

Stage 3 -- The information extracted during a series of fixations must be integrated into some overall representations.

Loftus reports research that supports the view that readers obtain what is referred to as "gist" information about the graphic within the first one in two eye fixations. Gist information is simply information that permits the readers to say what the graphic represents in some general way (i.e., a picture of a horse vs. a picture of a lion).

Next it appears that readers fixate on what are referred to as informative areas in the graphic. Such areas are defined by Mackworth as areas reported by subjects to be informative while Berlyne (1950) defines these

areas in terms of their complexity and novelty. (For example, subjects viewing a city skyline would find the outline of a very tall building to be informative as opposed, say, to the information provided by an area of clear sky).

The eye-track data provided by our videotapes seems to indicate that the addition of text to graphic presentations does not interfere with the acquisition of gist information from the graphic.

That is, our videotapes show that the first second or two of viewing text graphic combinations is spent in obtaining what we believe is gist information of the graphic. Readers then tend to refer to the text and to read for several seconds. They then refer periodically to the graphic, seemingly for edification of information presented in the text. It is our supposition that these referrals from text to graphic are made to high information areas that are either redundant of the information content of the text just read or which complement such information.

In order to determine the accuracy of the suppositions we have begun to design text-graphic materials using the Frederiksen scoring system described earlier to control for semantic content. We have also begun to develop an on-line computer system to provide higher quality eye-track data. This system (shown in Figure 6), developed by George McConkie of Cornell University, displays text and graphics on a cathode-ray tube.

-- Insert Figure 6 about here --

The reader, whose eye is shown at the upper left corner of the figure, is looking at the scene while a "Device" simultaneously observes the eye. The device includes detectors that acquire information about eye position. This information is converted to signals which are further converted into a form usable by the computer.

At this point the information may be displayed on a slave scope to keep the experimenter informed as to the reader's activity, it may be recorded on a disc or magnetic tape for future analysis or it can be returned to the cathode ray tube in such a way as to modify the display contingent upon certain behaviors of the reader. If, for example, a reader looks for a given number of seconds at a portion of the graphic, the computer could be programmed to enlarge that section, or add additional details, etc.

Although we have developed the programs for generating text and graphics for display purposes and have written the data collection and some data analysis programs, we are just this month about to conduct our first pilot study with this equipment.

The computer which is used in this eye-track research is also the solution to our data management problem, thus satisfying the fourth requirement for our research. The facility has been developed by George McConkie and has been generously made available for our use.

At this point I would like to present some of the very tentative findings of our earliest research.

Review of Research

Our first efforts were directed towards learning how differences in the form of semantic content would affect reading comprehension as measured by an assembly task. We developed text and illustrations that provided instructions for the assembly of a model loading cart or hand truck. At that time we did not have the precise measure of semantic content that I have earlier described, but several independent judgments were obtained to attempt to control for this variable.

We attempted to express the identical semantic content in the form of

text, illustrations of several types, and a combination of text and graphics. We expected that variations in form would result in variations in readers' comprehension, despite the fact that the content of the presentation would be the same in every case.

This expectation was based on the results of a series of pilot studies in which it was found that the different types of information contained in a discourse are (depending on the specific type of information) better communicated by text, graphics or as text-graphic combinations.

To date, these studies have been conducted to examine this issue, using variations in the form of text and several variations in the form of graphics. These variations have included variations in the amount of detail in the text, variations in the amount of detail in graphics (from line drawings to shaded line drawings), the use of symbols such as arrows or dotted lines, the use of exploded views, and so forth.

Although our analysis of these data is not yet complete, some findings do seem to appear consistently in each study. We find that, as indicated in our pilot studies, certain specific types of information appear to be better understood when presented in a particular form of text, graphics, or combinations of the two.

One type of information found in the directions used in these studies may be called spatial or locative information. That is, information concerning any or all of the three spatial dimensions. Locative information may specify the location and/or orientation of an object or features of an object within the three dimensional spatial matrix. Such information may take the form of a sentence such as "the pen is on the table" or may be as complex and information-filled as a topographic map of Europe.

We have consistently found that such information is more completely

and accurately understood when presented in graphic forms than in text when the semantic content of each form is equivalent. We have found that it is a characteristic of text that information must be written and read in a linear sequence. At the time of writing, the writer must impose an order on the content to be written. One thing must be mentioned and then another. The reader must generally follow the order of presentation.

When locative information is presented in the form of text it seems to inevitably create problems for the reader. For example, if we wish to describe the construction of a house of cards it is not necessary to specify which cards we are referring to. That is, identifying a particular card such as the "Queen of Hearts" or "Jack of Diamonds" vastly complicates the task, since the "house" may be built without such information. In other words, the use of text seems to free us to express locative information in a cumbersome way (i.e., "the card resting on the right edge of the card on the left") or to be overly specific.

We also have tentative evidence that sequential information, or information concerning the order in which events occur is more effectively communicated in text than in still illustrations.

We are attempting to test these findings and statements about the design of text and graphics that are found in the technical writers' literature. We hope, in time, to develop a model of the reading process that will predict how people will read and understand various forms of text and graphics and to build a body of research upon which the improved design of text and graphics can be based.

References

Berlyne, D. E. The influence of complexity and novelty in visual figures on orienting responses. Journal of Experimental Psychology, 1958, 55, 289-296.

Frederiksen, C. Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 1975, 7, 371-458.

Loftus, G. General software for an on-line eye-movement recording system. Behavior Research Methods and Instrumentation, 1975, Vol. 7(2), 201-204.

Loftus, G. R. and Bell, S. M. Two types of information in picture memory. Journal of Experimental Psychology: Human Learning and Memory, 1975, Vol. 104, No. 2, 103-113.

Mackworth, N. H. and Bruner, J. S. How adults and children search and recognize pictures. Human Development, 1970, 13, 149-177.

Mackworth, N. H. and Morand, A. J. The gaze selects informative details within pictures. Perception and Psychophysics, 1967, 2, 547-552.

McConkie, G. W. Studying retention from prose: A content structure approach. Paper presented at the Annual Meeting of the American Educational Research Association, April, 1972.

Stone, D. E. Comprehension of information in picture-text amalgams in procedural texts. Unpublished Doctoral Dissertation, Cornell University, 1977.

Symbol	Example	Definition
'	('RUN)	Elementary (primitive) concept.
:	(:BOAT)	Particularized object. On object list.
--REL->	--AGT-> <-ORD-- <-IDENT->	Unidirection or Bidirectional relationships.
[ ]	["P52]	Proposition.
"	"P52	Indicates that the number following is a proposition number.
∂	AGT∂TEM(PAST)	Operator. Followed by operator name.
⌗	(⌗3.2)	Indicates that the number following is simply a slot identifier. The number is the number of proposition, plus a digit following a decimal if there are more than one empty slots in the proposition
*		Location of a branchpoint.
⌠		Marks the beginning of a branch line. Assumed to associated with the last branchpoint.
+		Marks line as a continuation from the previous line.
,	[(:CAR), :BUS)]	Used to separate elements in a list.
.	(:OCEAN.FISH)	Concatenates segments into a single label. Used to permit multi-word labels.
( )	(:WORD)	Used to indicate which may be filled by concepts or arguments.
!T	(!T21.5)	Like ⌗, except that it is used exclusively in slots following temporal relations.
;		Marks line as comment line, to be ignored in network operations.

Legal Symbols in Column One:

P- followed by a number which is taken as a proposition number.

⌠, which marks line as a branch line.

+, which marks line as a continuation from the previous line.

(, which marks line as a line in the object list.

;, which marks line as a comment line.

Acceptable format is as follows:

Network line (begins with -P followed by a number, ∫ or +)

Proposition number must be followed by a single tab.

Spaces are not permitted; they are reserved for use by the computer in reformatting the network.

Object list line (begins with "(").

Begin with label by which the object is to be identified in network, followed by tab (number of tabs is not critical).

Spaces are not allowed.

Example:

(:BOAT.SAIL) ('BOAT)--DEF->( )--NUM->('3)
Refers to a definite set of 3 boats.)

Labels should not contain more than 12 characters. Longer labels will be truncated by certain programs.

Figure 1

The Back

Move the two columns so that they are parallel. They should be two block widths apart (be sure that the tabs at the end of each column point in the same direction). Connect the two columns by inserting four flat pieces between them. Slide each flat piece into place using the side grooves in the blocks. (All the flat pieces should have their smooth sides on the same side.) They must be flush with the ends of the columns without tabs.

Figure 2

The Back

p. 111 ( )--agtətem(present)-->('MOVE)-->obj--('COLUMNS)--DEF-->(THE)
+ --NUM-->($^{1}2$)

p. 112 (:COLUMNS)--agtətem(present)-->('ARE)--att-->('PARALLEL)

p. 113 (:COLUMNS)--agtətem(present)-->('ARE)--loc-->('APART)--+att-->
+('BLOCK.WIDTHS)--NUM-->('2)

p. 114 ( )--agtətem(present)-->('ARE)-->att-->('SURE)

p. 115 (:TABS)--DEF-->(THE)--locəp('AT)-->(:END)-->
+locəp('OF)-->(:COLUMN)--DEF-->(EACH)-->
+loc-->('POINT)-->obj-->(:DIRECTION)--att-->
+(SAME)--DEF-->(THE)

p. 116 ( )--agtətem(present)-->('CONNECT)-->obj-->(:COLUMNS)+--DEF-->
+(THE)--NUM--('2)

p. 117 ( )--agtətem(present)-->('INSERT)-->(:PIECES)--+att-->('FLAT)-->
+NUM--('4)--locəp('BETWEEN)+-->(:THEM)

p. 118 ( )--agtətem(present)-->('SLIDE)--obj-->(:PIECE)-+--att-->('FLAT)
+ --DEF-->(EACH)+--locəp('INTO)-->(:PLACE)

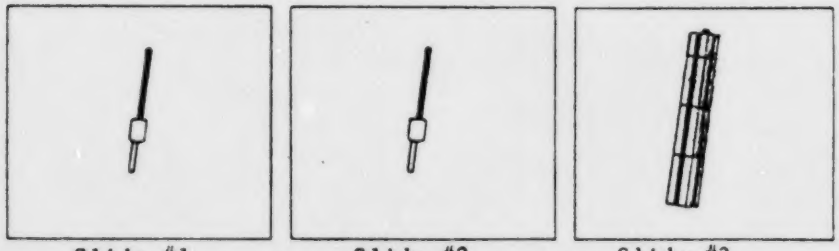
p. 119 ( )--agtətem(present)-->('USE)--obj-->(:GROOVES)+--att-->('SLIDE)
+ --locəp('IN)-->(:BLOCKS)+--DEF-->(THE)

p. 120 (:PIECES)--agtətem(present)-->('HAVE)--obj-->(:SIDES)--att--('SMOOTH)
+-loc---p('on)-->:SIDE)--+att-->(SAME)--DEF-->(THE)

p. 121 (:PIECES)--agtətem(present)--('BE)--att-->('FLUSH.WITH)-- obj-->
+(:ENDS)--DEF-->(THE)--locəp('OF)--obj-- (:COLUMNS)--att-->
+('WITHOUT.TABS)

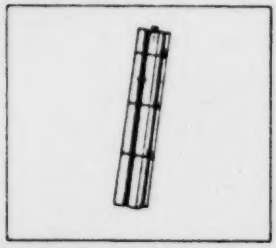
Figure 3

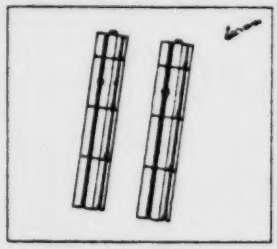
SLIDES OF ILLUSTRATION
(Not shown in accurate scale)

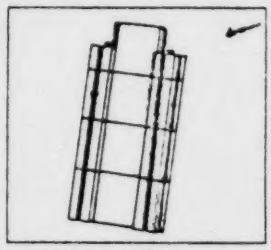

Slide #1

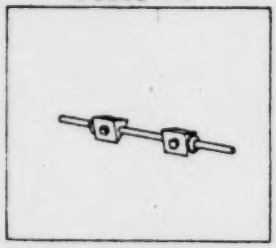
Slide #2

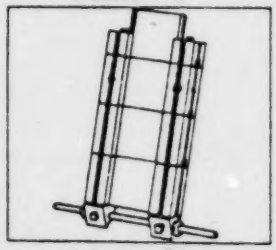
Slide #3

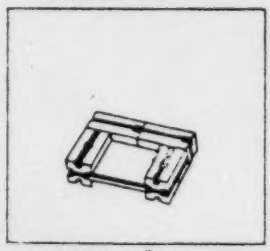

Slide #4

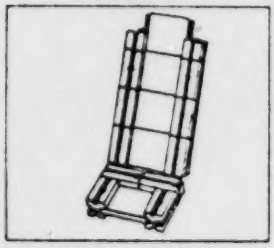

Slide #5

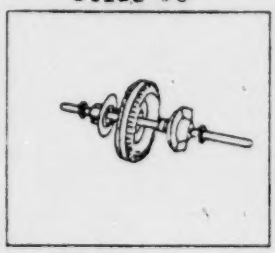

Slide #6


Slide #7


Slide #8


Slide #9


Slide #10


Slide #11

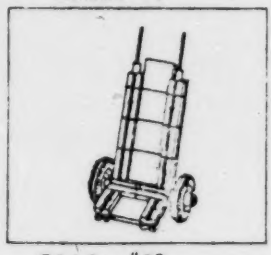

Slide #12

Figure 4

The Back

```
Slide #5

p. 111    -NOT PRESENT-

p. 112    (:COLUMNS)--agtətem(present)-->('ARE)--att-->('PARALLEL)

p. 113    (:COLUMNS)--agtətem(present)-->('ARE)--loc-->('APART)-- att-->
              +('BLOCK.WIDTHS)--NUM-->('2)

p. 114    -NOT PRESENT-

p. 115    (:TABS)--DEF-->('THE)--locəp('AT)-->(:END)
                              +--locəp('OF)-->(:COLUMN)--DEF-->(EACH)-->
                              +--loc-->('POINT)-->obj-->(:DIRECTIONS)--
                                   +att--> ('SAME)--DEF-->(THE)

p. 116    -NOT PRESENT-

Slide #6

p. 000    (:PIECES)--att-->('FLAT)--agtətem(present)-->(CONNECT)-+-obj-->
              +--NUM-->('FOUR)
          +(:COLUMNS)--DEF-->(THE)--NUM-->('2)

p. 117    -NOT PRESENT-

p. 118    -NOT PRESENT-

p. 119    -NOT PRESENT-

p. 120    -NOT PRESENT-


p. 121    (:PIECES)--agtətem(present)--('BE)--att-->('FLUSH.WITH)--
              +obj-->(:ENDS)--DEF-->(THE)--locəp('OF)
                   +--obj-->(:COLUMNS)--att--('WITHOUT.TABS)
```

PHYSICAL LAYOUT OF THE RESEARCH LABORATORY

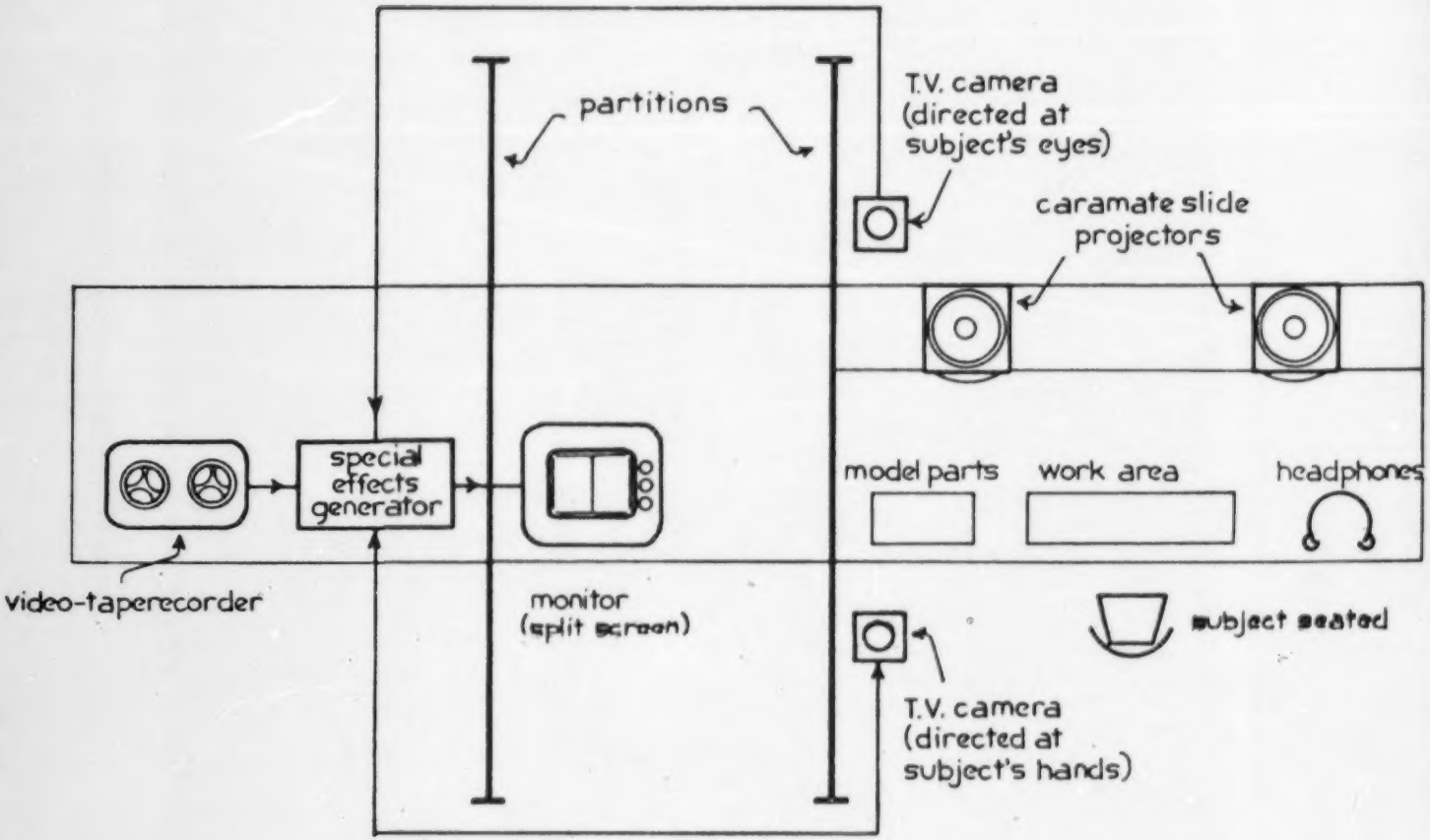
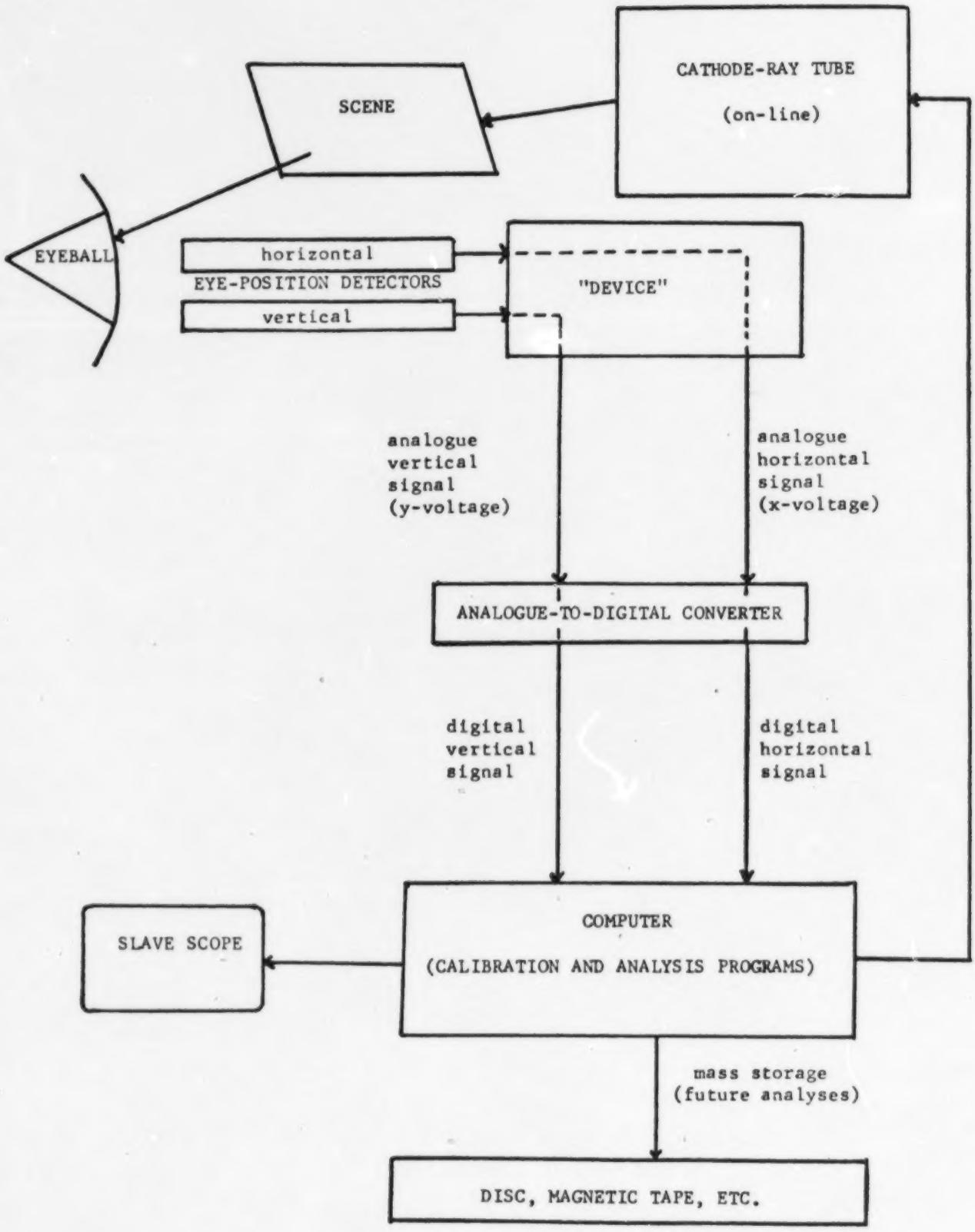


Figure 6



A general system for performing an experiment in which stimuli are controlled and eye movements are recorded on line.